

Slips of the tongue and language production. Edited by ANNE CUTLER. (*Linguistics* 19:7-8.) Berlin: Mouton, 1982. Pp. 293. DM 40.00

Reviewed by DONALD G. MACKEY and DANIEL KEMPLER, *UCLA**

This volume constitutes the first 'special topic' issue of *Linguistics*, published both as a book and as a double issue in the regular series. It comes in paperback, with a more affordable price tag than many edited volumes currently on the market. The contributions, either invited or selected from responses to a call for papers in the journal, include: Anne Cutler, 'Guest editorial: The reliability of speech error data' (561-82); Manfred Bierwisch, 'Linguistics and language error' (583-626); Brian Butterworth, 'Speech errors: Old data in search of new theories' (627-62); Andrew Crompton, 'Syllables and segments in speech production' (663-716); David Fay, 'Substitutions and splices: A study of sentence blends' (717-50); Jean Aitchison and Miron Straf, 'Lexical storage and retrieval: A developing skill?' (751-96); Paul Meara and Andrew W. Ellis, 'The psychological reality of deep and surface phonological representations' (797-804); Alan Garnham et al., 'Slips of the tongue in the London-Lund corpus of spontaneous conversation' (805-18); and Carol A. Fowler, 'Review of *Errors in linguistic performance*, ed. by Victoria Fromkin' (819-40).

The book tackles two basic questions of interest to philologists, phoneticians, phonologists, linguists, and psycholinguists working on the internal lexicon and sentence formation processes, as well as to psychologists with a Freudian bent. One is the question of how humans produce phonemes, words, and sentences—and how they acquire these abilities. The other concerns the special problems that arise in the use of speech errors as linguistic evidence. The answers offered here to these questions are sketchy and problematic; but the book should prove useful for seminars at the graduate and advanced undergraduate levels, and for anyone interested in speech errors and the detailed nature of the components and processes underlying speech production.

In over-all structure, the book resembles a sandwich with a salad on the side.

* We thank Victoria Fromkin for helpful comments on an earlier version of this review.

The outer layers of the sandwich are the articles by Cutler and by Garnham et al. (including Cutler). In her 'guest editorial', which introduces the book and establishes its themes, C regards existing collections of speech errors as flawed because they fail to consider the relative perceptibility of different types of errors. The argument goes as follows: even under the best of conditions, the most careful collectors are likely to misperceive or overlook some slips. And without the help of high-fidelity recording and painstaking, multiply-checked transcriptions, one cannot possibly detect, identify, recall, and record all the speech errors that one encounters in everyday life. As a consequence, 'more error' statements of the form 'More errors resemble X than Y', and 'no error' statements of the form 'No errors of type Z occur', are logically flawed—unless rather elaborate corrections for the relative detectability of X, Y, and Z are made.

To avoid such problems, C gives us two ways out. One is to make 'some error' statements rather than 'more/all/no error' statements. Unlike the latter type, which take the form 'More (all/no) errors resemble X (rather than Y)', the former take the form 'Some error(s) resemble(s) X.' C praises such statements as easy to make (requiring only a single uncontested example), and safe from perceptual and other sampling biases; she cites several examples within the book for possible emulation.

A second way out is to use a corpus of speech errors resembling the one collected by Garnham et al., which constitutes the other layer of the sandwich. These errors (191 in all) were obtained from Svartvik & Quirk's massive London-Lund corpus of spontaneous conversation (1980), and were cross-classified by type of error (in descending relative frequency: substitutions, anticipations, blends, omissions, other errors [sic], additions, perseverations, and exchanges), and by the size of the output units involved (words, segments, syllables, phrases, and other units—again in descending frequency). Because the errors were originally recorded on magnetic tape, the authors claim that the frequency estimates provided are unbiased by factors such as relative perceptibility.

If these claims are correct, the London-Lund corpus will become the focus of much future research; and C's perceptibility corrections represent an important methodological contribution, which will change forever the direction and nature of research into naturally occurring speech errors. However, there are important reasons for questioning these claims.

THE LONDON-LUND CORPUS: Does this corpus really help eliminate collector error? Previous collectors have consulted speakers about what they INTENDED to say, in order to (a) rule out possible misperceptions, and (b) determine the target utterance (i.e. the speaker's intention). These data enabled collectors to determine the category of an error and its possible cause, and to exclude otherwise irrelevant data—reflecting ignorance, false starts, intentional humor etc. But since the London-Lund corpus was obtained from verbatim transcripts without the possibility of speaker consultation, Garnham et al. must guess at these fundamental characteristics of the errors. For example, consider this error (257): *a series which ... are ... aimed at*. The authors classified this as a 'number agreement error', with putative target *is* instead of *are*. Perhaps, however, the speaker intended to say *series of things*, but omitted *of things*. Or perhaps the speaker thought *series* to be plural (like *scissors* or *pants*), and mistakenly inserted the article *a*. We will never know either the intended target or the true category of this or of many other errors in the London-Lund corpus.

Without independent confirmation, this of course makes suspect the relative frequency of the categories in this corpus.

Similar problems arise in the Fay and Aitchison–Straf corpuses published as appendixes to their articles. The Aitchison–Straf corpus consists of 680 putative word substitutions, obtained from readers of the *Sunday Times* in response to requests for errors of this type. How many of these solicited errors were mispronunciations rather than word substitutions? How many were caused by mishearing, ignorance, or intentional humor (on the part of either the perpetrator or the respondent)? In short, how many were really speech errors? We will never know for sure.

THE 'SOME ERRORS' APPROACH: Problems with this approach have been evident for some time (cf. MacKay 1973:787). Many factors can simultaneously influence any given error in everyday speech; most of these are neither known by, nor under the control of, the observer. Thus it is usually impossible to tell whether a single error or small set of selected errors reflects (a) the hypothesized factor, (b) some unknown and uncontrolled factor(s), or (c) both. Only when the hypothesized factor is found to have a systematic effect in large numbers of errors can firm conclusions be drawn. Such conclusions must be expressed as 'more errors than would be expected by chance' statements—which (unlike some/all/no error statements) require statistical tests of the sort that are implicitly avoided throughout most of this book.

THE PERCEPTIBILITY PROBLEM AND ITS PROPOSED SOLUTION: C claims that many slips go undetected, some more than others. The evidence for these claims is Tent & Clark's 1980 experiment on the detection of 'artificially constructed' slips in auditorily presented sentences. White noise of equivalent loudness masked the input. The subjects were linguistically unsophisticated, and had no prior experience or training in collecting errors. Indeed, they didn't even expect errors, and thought the experiment was designed to investigate the perception of everyday speech under noisy conditions. After being instructed to transcribe each sentence exactly as they heard it, they were said to 'misperceive' a slip if they supplied the original word, and to 'perceive' it if they supplied another word, left a blank, or correctly transcribed the actual phonetic form!

Can this experiment be said to simulate the conditions under which errors are normally collected, and provide a basis for correcting existing collections? Surely not. Everyday conversations rarely occur under circumstances even remotely resembling those of this experiment. True collectors are linguistically sophisticated; they expect errors, and have extensive experience in perceiving and transcribing them. They have the added benefits of discourse context, pragmatic situation, and face-to-face contact with the speaker. After an error, they can usually interact with the speaker to determine the target and verify what was said. Even when the collector misses an error, the speaker often saves the situation by making a correction, usually prefaced by telltale markers such as *I mean* or *um* or *er*. Sometimes, of course, the speaker is also the collector, who then has kinesthetic and bone-conducted feedback in addition to other cues for detecting the error. With regard to the perceptibility issue, true collectors are much less likely to miss or misperceive errors than C suggests. And even if she is NOT crying wolf in the absence of real wolves, she has certainly misadvised us as regards what to do about them.

Consider now the meat of the volume, which comes sandwiched between the methodological warnings of the first article and the raw data of Garnham et al. These central chapters are diverse in framework, methods, theory, and depth of detail. This heterogeneity seems unsatisfactory for an edited book—where the editor chooses the authors, and bullies them into contributing something that fits an integrated whole; it surely reflects the origin of this collection as a journal—where the authors are unselected and unconstrained, and the editor delegates the decision on acceptance to referees. Publishers take note: journals and books are fundamentally different beasts.

The contribution by Bierwisch is really two articles. Bierwisch I is an English translation of a 1970 paper (published originally in German) which discusses how various types of speech errors bear on models of language structure and performance. It concludes that speech errors (a) are grammatically principled, (b) result from interferences with underlying processing mechanisms,

and (c) fall into a relatively small number of basic categories (blends, selection errors, and sequential errors). Bierwisch II is a postscript (several pages longer than the original) which expands and updates Bierwisch I by contrasting speech errors with errors in action. It suggests that speech errors share some structural features with errors in other domains, but also have unique features. It cautions AGAINST using speech error for drawing conclusions about grammatical structure; and it suggests that, although constrained by language-specific structures, speech errors are caused by general mechanisms which are not restricted to language. Both contributions are observational, speculative, and tentative, but thoughtful. More than any other contributions to the present volume, they use speech error data to address crucial points in syntactic theory, e.g. the controversy over the lexicalist vs. transformational view of the lexical store.

The remaining chapters are more concerned with error data per se than with theoretical issues, and seem bent on 'accounting for' the errors by constructing special 'speech error theories'. All devote an inordinate amount of space to defending definitions or analyses which lack behavioral or theoretical rationale.

The chapters by Crompton and by Butterworth present us with no new data, but attempt to review and revise existing theories of speech production derived from speech error data. Both these papers address important issues (e.g. sequential vs. parallel processing), but exhibit a notable lack of concern for parsimonious explanations of how normal, error-free output is achieved, or for other data bearing on this issue. Crompton attempts to delineate the interaction of the structural units of speech production (features, phones, syllable constituents, and syllables) in the construction of articulatory programs. He also addresses the (at last) fashionable topic of the status and structure of syllables in phonology; but he seems unaware of extensive work on this subject undertaken in America (e.g. MacKay 1973, 1982; Treiman et al. 1982). Butterworth's critiques of Freud's, Fromkin's, and Garrett's theories are especially well argued, and his graphic illustrations are clear and helpful for anyone unfamiliar with speech production models; but it should be noted that many of the assumptions and proposals that he criticizes in the model of Fromkin 1971 have already been abandoned by her in print (Fromkin 1980).

Two papers (Fay and Meara & Ellis) are devoted to categorizing new but relatively small bodies of error data. Both are sophisticated but descriptive, attempting only to 'make sense of the data'. Fay's analysis of 48 blends indicates that substitution blends have different characteristics from splice blends, and that duplication errors may or may not be blends. Meara & Ellis contribute nine Welsh spoonerisms provided by native speakers. Although the data base is small and their analyses brief, this appears to be the first published account of speech errors in a mutating language, where word-initial consonants are determined in part by their syntactic environment.

Aitchison & Straf's chapter is the only one which focuses on normal functioning. Their question is whether factors influencing the perception, storage, and retrieval of words differ for children vs. adults. To answer this question, they examined 680 malapropisms, e.g. substitution of *genitals* for *gentiles*. They divided these errors into two speaker groups (under vs. over 13 years old), which they compared using a statistically elegant multivariate analysis. One example of the results is that intended and substituted words tended to share the features assessed (initial consonant, final consonant, number of syllables, vowel with primary stress, and over-all stress pattern). However, child malapropisms shared the same number of syllables relatively more often, whereas adult malapropisms shared the same initial consonant relatively more often. It is a shame that more reliable and age-specific data (e.g. broken down into smaller age-groups) were not available for this study: the age comparisons might have revealed quantitative developmental progressions of the type investigated elsewhere in language acquisition research. However, because of data limitations, the authors could conclude only that 'child and adult lexical storage and retrieval should be studied separately'. This conclusion may shock the many researchers who are doing an excellent job of comparing the acquisition and use of lexical knowledge by children vs. adults, and may comfort the many others who have studied the two populations separately over the past hundred years.

We come at last to the salad, Fowler's 21-page review of Fromkin 1980. Since Fowler is in the speech production field, but is not a speech error person, she provides a valuable perspective on the contribution of speech errors to the

field at large. However, we had mixed reactions to its position on the menu: last rather than first. It has a valuable introductory function: to ground the present volume in a well-established tradition of edited volumes concerned with speech errors. Fowler gives positive comments on some of the methodological and theoretical directions in Fromkin 1980 which are sorely lacking in the present volume; if it had come first (and been taken seriously), it might have greatly influenced the remainder of the book.

Where do we go from here? Our review of C's book suggests several directions for the future, the first of which is to develop a truly theoretical orientation. One searches in vain in the book for theoretically based predictions, or even hypotheses; what theoretical statements are present are ad hoc, metaphorical, and unconvincing. Many are pseudo-explanations which invoke homunculi such as bookkeepers, librarians, text generators, programmers, scan-copiers, editors, and incorporators (in addition to their many other duties, incorporators sequence words, presumably by consulting a grammarian homunculus). Other attempts at explanation resemble old wine in new bottles. We cannot be content to apply labels such as 'misperception' to an error, or to relabel Freudian slips in new ways. Speech errors are clues to a complex underlying system, windows to the mind. Our goal is to infer the nature of that underlying system from the way it breaks down. Our question is how and precisely WHAT processes break down so as to produce the observed errors; and no labeling or relabeling will suffice to answer this question.

The second direction for the future is to become less insular. One is astounded, e.g., at Cutler's claim that attempts to demonstrate that error rates rise with rate of speech have all met with failure. The claim is not only false (cf. MacKay 1971), but flies in the face of a very general and well-established phenomenon (speed-accuracy trade-off) which is replicated throughout the field of psychology (cf. MacKay 1982; Meyer et al. 1982). We cannot afford to overlook research within our own and related domains. It will not do for each new investigator to pretend to re-invent our long established fields; nor will it do to view speech errors as a separate field, with its own special methodology and phenomena. Speech errors are only one brand of data for addressing linguistic and psycholinguistic issues (cf. Baars et al. 1975 for integration of experimental and descriptive methods). Only by solidifying our conclusions with every available method, and by developing theories which are general and viable, will we begin to converge on the truth.

REFERENCES

- BAARS, B.; M. MOTELY; and D. G. MACKAY. 1975. Output editing for lexical status from artificially elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior* 14.382-91.
- FROMKIN, VICTORIA. 1971. The non-anomalous structure of anomalous utterances. *Lg.* 47.27-52.
- . 1980 (ed.) *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand*. New York: Academic Press.
- MACKAY, DONALD G. 1971. Stress pre-entry in motor systems. *American Journal of Psychology* 84.35-51.

- . 1973. Complexity in output systems: Evidence from behavioral hybrids. *American Journal of Psychology* 86.785-806.
- . 1982. The problems of flexibility, fluency, and speed-accuracy trade-off in skilled behavior. *Psychological Review* 89.483-506.
- MEYER, D. E.; J. E. K. SMITH; and C. E. WRIGHT. 1982. Models for the speed and accuracy of aimed movements. *Psychological Review* 89.449-82.
- SVARTIK, JAN, and RANDOLPH QUIRK. 1980. *A corpus of English conversation*. Lund: Gleerup.
- TENT, J., and J. E. CLARK. 1980. An experimental investigation into the perception of slips of the tongue. *Journal of Phonetics* 8.317-25.
- TREIMAN, R., et al. 1982. Effects of syllable structure on adults' phoneme monitoring performance. (Speech Research Laboratory Program, Report no. 8.) Bloomington; Indiana University.

[Received 4 May 1983.]